Advances in Computer, Signals and Systems (2024)
Clausius Scientific Press, Canada
DOI: 10.23977/acss.2024.080402
ISSN 2371-8838 Vol. 8 Num. 4

A Safety Helmet Detection Method Using Adjusted YOLOv8

Minghao Yang[1], Feng Rao[2], Lili Wang[2,*], Mingsong Bai[1], Xiqing Zang[1]

[1]Beijing Research Institute of Automation for Machinery Industry, Beijing, 100120, China
[2]RIAMB (Beijing) Technology Development Co., Ltd., Beijing, 100120, China
[*]Corresponding author

Keywords: Safety helmet detection, YOLOv8, DenseNet, safety helmet wearing dataset

Abstract: Safety production is of paramount importance in protecting the safety, health of workers and assets. Safety helmets play a crucial role across various industries, directly impacting the wearer's life safety. In response to the prevalent issue of many workers not wearing safety helmets, coupled with high cost and risks associated with manual safety helmet detection, current automated methods are difficult to detect safety helmet usage at a large scale, complex on-site environments. This paper proposes a safety helmet detection method based on adjusting YOLOv8. Adjustments to the backbone network of YOLOv8 were replaced by DenseNet121 and appropriate data augmentation methods were designed. This method achieved an accuracy of 96.81% in the Safety Helmet Wearing Dataset. Compared to the original YOLO v8 algorithm, it achieved a 0.74% performance improvement. Our method enhances the accuracy of safety helmet detection, provided important technical support to ensure production safety.

1. Introduction

Safety production is of paramount importance in protecting the safety, health, of workers and national assets, while also serving as a cornerstone for the development of social productivity. Ensuring proper safety measures is of paramount importance. Safety helmets, as a form of personal protective equipment, play a crucial role across various industries, directly impacting the wearer's life safety. Designed to protect the head, they mitigate injuries from unforeseen accidents, offering functions such as impact resistance, abrasion prevention, puncture resistance, and electrical insulation. In industries such as construction, mining, power, and transportation, wearing safety helmets is mandatory.

Safety helmets are essential protective tools for workers, yet many choose not to wear them due to discomfort, posing a threat to their safety. Therefore, real-time monitoring of whether workers are correctly wearing safety helmets is crucial. The hazardous and complex working environment on construction sites is not suitable for comprehensive manual monitoring. Moreover, compared to traditional manual inspections, automated detection systems can identify instances of non-compliance with helmet-wearing regulations more quickly and accurately, saving on labor costs.

Existing algorithms for detecting the usage of safety helmets can be broadly categorized into two main types: the traditional machine learning-based object detection algorithms and the deep learning-based object detection algorithms. Traditional algorithms primarily extract features for safety helmet detection that are low-level and manually selected, such as color and shape [1]. Including the helmet recognition method based on edge detection [2], SVM and skin color detection [3], the method with modified deformable part model by a combination of histogram of block-based local binary pattern [4], the method used Haar-like features to detect faces and used an edge detection algorithm to find helmet contour features [5] and the color-based hybrid descriptor with H-SVM method are utilized for helmet identification [6]. Due to their simple algorithmic structure, traditional algorithms require less computational resources and exhibit faster detection speeds. However, in complex project site environments, the detection accuracy of traditional methods for safety helmet detection is relatively low, resulting in significant disparities from the management requirements in actual field settings. Deep learning-based methods offers high precision and flexibility. The majority of deep learning-based methods focus on object detection of safety helmets as a basis for determining whether workers are wearing them or not. Additionally, these methods can handle large volumes of image or video data, making them suitable for monitoring multiple construction sites or work areas. Such scalable applications contribute to enhancing overall safety. There is immense potential in applying these techniques to safety helmet detection, offering new avenues for improving the production safety of on-site workers. For example, the helmet detection methods based on SSD [7-8], RCNN series [9-10] and YOLO series [11-13]. Although these deep learning-based methods have achieved some success in the field of safety helmet detection, there are still issues such as insufficient accuracy and poor performance in complex scenarios. Therefore, to address these concerns, we propose an adjusted YOLOv8-based safety helmet detection approach aimed at enhancing detection accuracy and adaptability to better meet practical application requirements.

2. Methods

Safety helmet detection technology is constrained by various objective factors. Deep learning algorithms need to balance accuracy with real-time performance. Additionally, considering the hardware costs for on-site deployment, it's essential for deep learning algorithms to have low computational and memory requirements, enabling efficient operation on resource-constrained devices.

To achieve real-time, high-precision identification of safety helmets in various on-site scenarios, we propose a method based on adjusted YOLOv8 algorithm [14] for object detection of safety helmets. Our work involves replacing the backbone network of the original YOLOv8 algorithm and designing appropriate data augmentation methods, with other adjustments about training strategies.

2.1 Original Model

YOLOv8 was developed by Ultralytics, stands as one of the current state-of-the-art models supporting multiple visual tasks. It builds upon previous iterations of the YOLO series, incorporating improvements and new features, resulting in enhanced performance. Compared to YOLOv5 [15], YOLOv8 demonstrates a significant improvement in accuracy on the MS COCO dataset [16], albeit with an increase in parameter count. YOLOv8 achieves higher accuracy while also maintaining model inference speed, ensuring stronger real-time capabilities. It offers various model scales, including YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large), and YOLOv8x (extra large), each optimized with lower computational and memory requirements. This enables flexible selection of model scales based on specific needs and efficient operation on resource-constrained devices. Consequently, YOLOv8 enables real-time object detection on lower specification devices without significant hardware investment, enhancing device compatibility and flexibility for detection.

The structure of YOLOv8 consists of three main components: the input, the backbone network, and the head (Figure 1). The backbone consists of convolutional module (ConvModule), cross stage partial module (C2f) and spatial pyramid pooling fusion module (SPPF). The main idea of YOLOv8 algorithm is partitioning the input feature map into multiple scales of grids, with each grid responsible for detecting targets in specific regions and generating prediction boxes for target detection by adjusting prior boxes associated with the grids. This concept continues the core philosophy of the YOLO series, which transforms the object detection problem into a single neural network prediction problem, achieving end-to-end processing of the entire image. YOLOv8 further introduces methods such as scale feature extraction and adaptive training strategies to achieve efficient, accurate, and versatile object detection results.

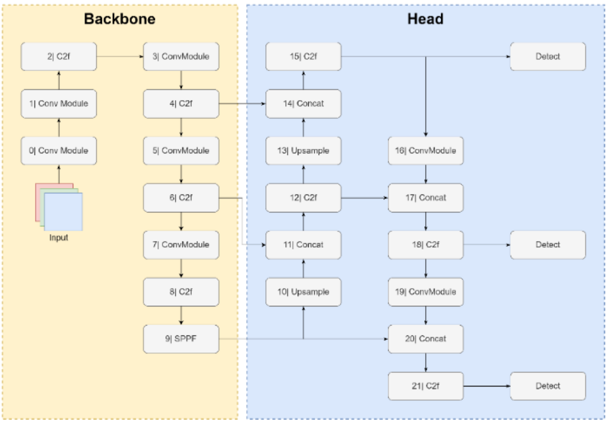

Figure 1: Network architecture of YOLOv8

2.2 Adjustments to the backbone

Although attention-based visual models have achieved tremendous success on many public datasets, their extensive data requirements and hardware demands make it difficult to achieve optimal recognition performance on smaller datasets. For these types of visual tasks, machine learning methods based on convolutional neural networks remain the primary approach for visual object recognition. In terms of balancing model recognition accuracy and device-friendliness, DenseNet [17] models stand out as exemplary.

The DenseNet network architecture was proposed by Cornell University, Tsinghua University and Facebook AI Research. Researchers introduced shorter connections between layers near the input and output layers of convolutional networks, enabling deeper, more accurate, and more efficient training. This concept replaces the traditional feedforward neural network connections with dense connections, utilizing the algorithm's feature maps as inputs for all subsequent layers. Unlike previous deep convolutional neural networks, DenseNet does not rely on extremely deep or wide architectures to extract representation capabilities. Instead, it leverages feature reuse to harness the network's potential. Whilst following a simple connectivity rule, DenseNets naturally integrate the properties of identity mappings, deep supervision, and diversified depth. Despite adding more connections, DenseNet only increases a small fraction of feature maps while keeping the rest unchanged. Therefore, compared to other architectures, DenseNet has fewer parameters and excels in computational and memory efficiency. Its unique layer-wise connections also alleviate the problem of gradient vanishing during neural network training, encourage feature reuse, and strengthen feature propagation.

With these innovations, DenseNet has significantly improved the performance of convolutional neural networks across various recognition benchmarks. Therefore, in this paper, we replaced the backbone of the original YOLOv8 network with the DenseNet121 network architecture (Figure 2), enabling the object detection model to achieve better results in safety helmet detection.

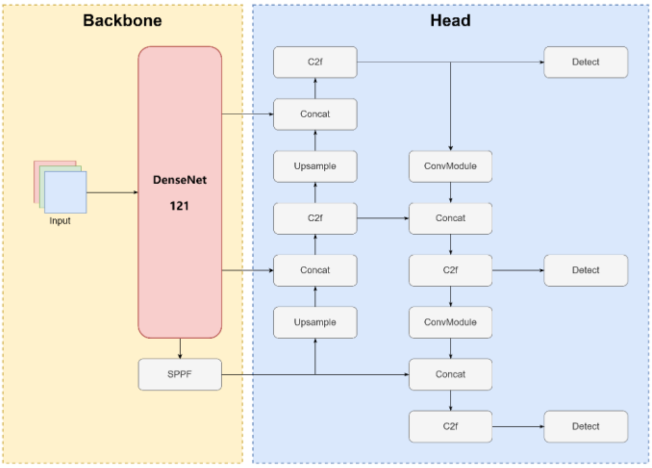

Figure 2: Network architecture of adjusted YOLOv8

2.3 Assessment Indices of the Drought

During the training process of deep convolutional neural network, if the training data is limited and singular, the model may only learn a limited data distribution, which could lead to overfitting on the training set, resulting in poor performance on the test set, low generalization ability, and susceptibility to noise and interference. In this paper, some data augmentation steps were taken when training deep neural networks for safety helmet detection. These steps aim to prepare training data to be more representative, diverse, and informative, thereby enhancing the model's performance and robustness.

3. Experiments and Results

Data plays a crucial role in the training process of deep neural networks. Safety Helmet Wearing Dataset (SHWD) provide the dataset used for both safety helmet wearing and human head detection. It includes 7581 images with 9044 human safety helmet wearing objects (positive) and 111514 normal head objects (not wearing or negative). In this paper, the Mean Average Precision 50 (mAP50) on SHWD was chosen for evaluating the performance of the models. The dataset is randomly divided into training and testing sets in a ratio of 80%-20%.

Data augmentation methods were tailored for this dataset. Partial training hyperparameters were outlined in the Table 1.

Table 1: Hyperparameters of adjusted YOLOv8

Hyperparameters	Value
epochs	200
imgsz	1088
optimizer	auto
close_mosaic	10
cos_lr	True
mixup	0.1
degrees	15
shear	15
dropout	0.1

According to the model training settings in Table 1, the performance of different models is shown in the Table 2.

Table 2: Performance of different models on SHWD

Algorithm	mAP
YOLOv4 [18]	91.4*
Faster R-CNN [19]	91.9*
YOLOv5 [12]	94.2*
YOLOv8 [13]	94.36
YOLOv5 [15]	95.31
YOLOv5 [15]	95.77
YOLOv8 [14] with default augmentation	96.07
YOLOv8 [14] with tailored augmentation	96.22 (+0.15)
Ours	**96.81 (+0.74)**

The accuracy and loss changes during the training process are shown in Figure 3 and Figure 4, The model using DenseNet as the backbone network has lower training and testing losses than the original YOLOv8 algorithm.

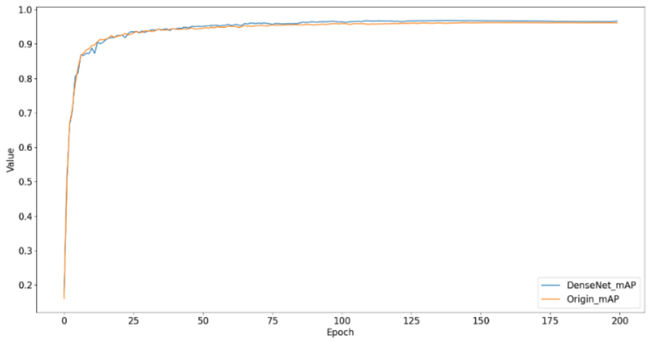

Figure 3: Changes in mAP with Epoch in different model

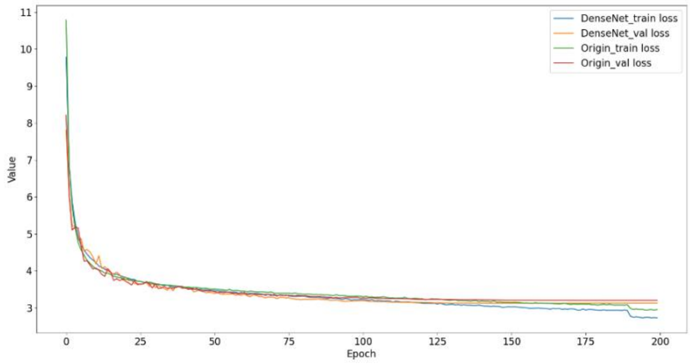

Figure 4: Changes in loss with Epoch in different model

Under the condition of Intel Core i9-12490F CPU and NVIDIA GeForce RTX 3080Ti GPU. The adjusted YOLOv8 safety helmet detection method achieved an accuracy of 96.81% in the SHWD dataset. Compared to the original YOLO v8 algorithm, our method achieved a 0.74% performance improvement.

4. Conclusions

To address the high cost and risks associated with manual safety helmet detection, as well as the inadequate accuracy and poor adaptability under complex on-site environments of existing automated methods, we propose an adjusted YOLOv8 safety helmet detection algorithm. By replacing the original YOLOv8 backbone with DenseNet121 and designing appropriate data augmentation techniques, our improved YOLOv8 achieved a 96.81% accuracy on the SHWD dataset, representing a performance increase of 0.74% compared to the original YOLOv8 algorithm. Our method enhances the accuracy of safety helmet detection and provided important technical support to ensure production safety.

Acknowledgements

The helmet images used in the manuscript were taken from the open-source Safety Helmet Wearing Dataset (https://github.com/njvisionpower/Safety-Helmet-Wearing-Dataset).

References

[1] Lihong Wei, Panpan Liu, Haihui Ren and Dong Xiao. (2024) Research on helmet wearing detection method based on deep learning. Scientific reports, 14(1): 7010.

[2] Kishor Shrestha, Pramen P. Shrestha, Dinesh Bajracharya, and Evangelos A. Yfantis. (2015) Hard-hat detection for construction safety visualization [J]. Journal of Construction Engineering, 2015(1): 1-8.

[3] Xiaohui Liu and Xining Ye. (2014) Skin color detection and Hu moments in helmet recognition research. Journal of East China University of Science and Technology, 3: 365-370.

[4] Junsu Jia, Qingjie Bao and Huiming Tang. (2016) Method for detecting safety helmet based on deformable part model. Application Research of Computers, 33(3): 953-956.

[5] Huang S, Huang J, Kong Y. Attention Guided YOLOv3 for Wearing Safety Helmet Detection[C]//ICCIP 2020: 2020 the 6th International Conference on Communication and Information Processing. 2020.DOI:10.1145/3442555.3442566.

[6] Xitian Long, Wenpeng Cui and Zhe Zheng. (2019) Safety helmet wearing detection based on deep learning. IEEE 3rd information technology, networking, electronic and automation control conference (ITNEC). IEEE, 2019: 2495-2499.

[7] Guang Han, Mengcheng Zhu, Xuechen Zhao and Hua Gao. (2021) Method based on the cross-layer attention mechanism and multiscale perception for safety helmet-wearing detection. Computers and Electrical Engineering, 95: 107458.

[8] Qi Fang, Heng Li, Xiaochun Luo, Lieyun Ding, Hanbin Luo, Timothy M. Rose and Wangpeng An. (2018) Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. Automation in construction, 85: 1-9.

[9] Mingyuan Zhang, Zhiying Cao, Xuefeng Zhao and Zhen Yang. (2019) On the identification of the safety helmet wearing manners for the construction company workers based on the deep learning theory. Journal of Safety and Environment, 19(02): 535-541.

[10] Yijia Zhang, Fusu Xiao and Zheming Lu. (2022) Helmet wearing state detection based on improved YOLOv5s. Sensors, 22(24): 9843.

[11] Yuejing Qian and Bo Wang. (2023) A new method for safety helmet detection based on convolutional neural network. PLoS one, 18(10): e0292970.

[12] Bingyan Lin. (2024) Safety Helmet Detection Based on Improved YOLOv8. IEEE Access.

[13] Ultralytics. (2023) Yolov8. https://github.com/ultralytics/ultralytics

[14] Ultralytics. (2020) Yolov5. https://github.com/ultralytics/yolov5

[15] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár and C. Lawrence Zitnick. (2014) Microsoft coco: Common objects in context. ECCV 2014, 2014: 740-755.

[16] Gao Huang, Zhuang Liu, Laurens van der Maaten and Kilian Q. Weinberger. (2017) Densely connected convolutional networks. CVPR 2017, 2017: 4700-4708.

[17] Benyang Deng, Xiaochun Lei and Miao Ye. (2020) Safety helmet detection method based on YOLO v4. CIS 2020, 2020: 155-158.

[18] Songbo Chen, Wenhu Tang, Tianyao Li, Huling Zhu, Ye Ouyang and Wenbo Wang. (2020) Detection of safety helmet wearing based on improved faster R-CNN. IJCNN 2020, 2020: 1-7.

[19] Zongzhe Lyu, Hui Xu, Xiao Yang, Yong Wang and Weijian Wang. (2023) Small object detection algorithm of YOLOv5 for safety helmet. Journal of Computer Applications, 43(6): 1943.